REMARKS

Applicants gratefully acknowledge the Examiner's determination that Claims 3-10 and 14-21 are drawn to allowable subject matter, as well as the Examiner's acceptance of the drawings submitted January 8, 2001.

Claims 1-21 are currently pending in the application. Claim 1 is currently amended in response to an objection raised by the Examiner by changing "means" to "engine" in line 7 to maintain consistency with the term "engine" in line 6 of the claim, thus correcting the informality. No new matter has been added.

The Claimed Invention

The claimed invention provides an extension to the HyperText Markup Language (HTML) allowing a user to employ context-sensitive audio commands to tell a browser what to present and what options are available for interaction with an application for which audio commands have been enabled. The claimed invention enables voice commands needed by an application, registers such commands with a speech engine, and provides an audio context for page-scope commands by adding a context option to make the page more flexible and usable. The invention thus enables a browser to respond to visual or verbal commands, or a combination thereof, by identifying what action will be taken based on the commands.

According to the prior art, applications, browsers, and speech engines are tightly linked together in a manner that prevents one application from working with multiple browsers or speech engines. As a result, current implementations have devices that will read aloud the words on a page but which require input to be entered either by keyboard or by an elaborate method such as where a user must proceed letter-by-letter using code words for letters of the alphabet, like "Alpha" for "A."

It is an object of the claimed invention to allow applications to register specific commands that will cause a browser to take an action based on the current audio context of the browser. It is a further object of the claimed invention to have a browser take an

action based on current audio context and a word or words currently being spoken by a user. It is yet another object of the claimed invention to allow one application to work with multiple browsers and speech engines.

The claimed invention provides a generic way of encoding information needed by an application to register voice commands and enable the speech engine. This is done by introducing new HTML statements with the keyword META_VERBALCMD, which list the recognized/registered speech commands and what each one will do. This applies to commands that affect a whole PAGE in scope, like the "help" or "refresh" command. No matter where a user is on the page or what the user is doing, these commands work the same and issue the same URL command to the user just as if the user had physically clicked on the HELP or REFRESH buttons on the screen.

The claimed invention further provides a sense of audio context. The context of a page changes as the audio presentation of the page progresses. The claimed invention adds the ability to alter the action based on the current audio context by adding the CONTEXT option to the META_VERBALCMD statements.

To take one possible example *inter alia*, the application may be a trip planner installed in an automobile and may be enabled to speak directions while displaying a map. A spoken command such as "repeat" may be employed to cause the application to speak the whole page of directions from the from the beginning. According to the claimed invention, however, it is possible to specify CONTEXT=OPTIONAL so that the browser may provide the application with a context to enable the application to tailor its response to the spoken command "repeat." Thus, if the user is listening to a direction at the time he or she speaks the command "repeat," the application would apply the command to the context and repeat the particular direction. If, however, the user is not listening to data from the application at the time she or she speaks the command "repeat" (*i.e.*, there is no current CONTEXT), the application would apply the command in the absence of context and speak the whole page of directions from the beginning.

Some spoken commands may be specified as CONTEXT="REQUIRED" instead of CONTEXT=OPTIONAL. To take one example *inter alia*, a person may be reviewing email in an audio mode while driving. While an email application is reading aloud topic of an email message or the name of the sender, a command such as "open" spoken by the user may cause the application to open and read aloud the contents of the message. According to the claimed invention, the performance of such an application could be improved by specifying CONTEXT="REQUIRED" to instruct the browser to recognize the spoken word "open" as a command only when there is an appropriate context recognized by the application at the time the word is spoken. If no such context is present when the word "open" is spoken, the word will not be recognized as a command. Thus, by way of example and not limitation, a user arriving at a rest stop may speak the command "stop reading" to stop reviewing email. Such user may then tell passengers, "You can open the door now and get out," without causing the email application to interpret the word "open" as a command to open an email message. This would occur because of the absence of an appropriate CONTEXT under circumstances in which CONTEXT="REQUIRED" has been specified.

Claims 1, 2, 11, 12, and 13 have been rejected under 35 U.S.C. § 102(b) as anticipated by U.S. Patent No. 5, 732,216 to Logan et al. Applicants respectfully traverse on the basis that Claims 1, 2, 11, 12, and 13 are not anticipated by Logan et al., as discussed below. Among other considerations, the "context based" features of the claimed invention enable commands to be registered at different times during a documents-audible presentation, and to permit commands to have different meanings at different times depending on the context. The disclosure of Logan et al. does not address context based commands in any way.

Claim 1. The Examiner has found that "[r]egarding independent claim 1, *Logan et al.* discloses a system for controlling an audio controller" and that the invention disclosed by Logan is equivalent to the claimed "system for providing context based verbal commands to a multi-modal browser." Applicants respectfully traverse.

Independent Claim 1 of the claimed invention, as amended above, provides as follows:

> A system for providing context based verbal commands to a multi-modal browser, comprising:
>
> a context-based audio queue ordered based on contents of a page being audibly read by the multi-modal browser to a user;
>
> a store for storing a current context of the audio queue; and
>
> a speech recognition engine for recognizing and registering voice commands, wherein said speech recognition engine compares a current audio context with the context associated with a voice command and causes the browser to perform an action based on the comparison.

(Claim 1, lines 1-9) Claim 1 thus registers audio commands obtained from the input markup and allows users to speak such commands to bring about the action of the renderer. The commands thus registered are dynamic in nature and need not be the same for every page, a feature not disclosed or anticipated by Logan et al. Even though the specification of Logan et al. mentions the use for audio commands for navigating the system, there does not appear to be anything to indicate that Logan et al. ever recognized problems relating to how to get a browser to take an action based on the current audio context of the browser. By contrast, Claim 1 is directed to "[a] system for providing context based verbal commands to a multi-modal browser," which is not accomplished or discussed by Logan et al. Nor does there appear to be anything in the disclosure of Logan et al. to anticipate "registering voice commands, wherein said speech recognition engine compares a current audio context with the context associated with a voice command and causes the browser to perform an action based on the comparison," as in Claim 1. While Logan et al. does appear to contemplate registration of commands, such commands appear to be fixed in nature, supporting only a standard set of navigation keywords designed to supplement convention automobile radio, tape of CD controls, and not context-sensitive as in Claim 1:

> The ability to navigate the program using only audio prompts and/or small number of buttons for a user interface make the playback system which utilizes these features of the invention particularly attractive for use by automobile drivers, who can select their program content much more effectively and with less drive distraction than currently possible with a conventional automobile radio, tape or CD player.

(Logan et al., column 35, lines 48-55) Applicants respectfully submit that the Examiner's finding that Claim 1 is anticipated by Logan et al. is based on a misapprehension of the reference, the claimed invention, or both.

In finding Claim 1 to be anticipated by Logan et al., the Examiner has relied extensively on Figure 5 from the disclosure of Logan et al.:

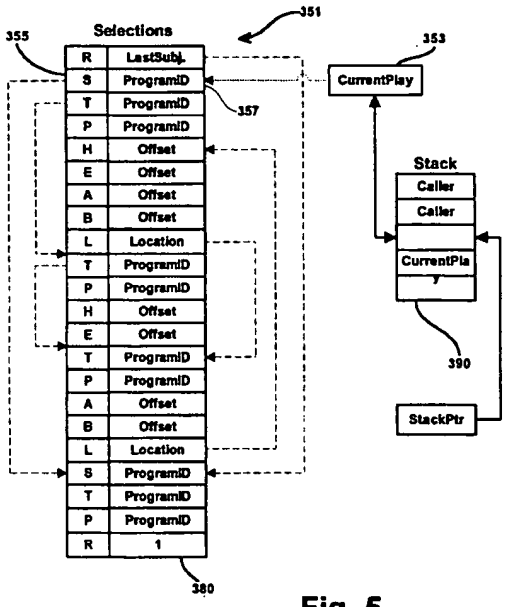


Fig. 5

(Logan et al., Figure 5) Nothing in Figure 5 of Logan et al. refers to a "multi-modal browser" and, because Figure 5 makes no provision for context sensitivity, there is nothing to anticipate a "context-based audio queue ordered based on contents of a page being audibly read by the multi-modal browser to a user," "a store for storing a current context of the audio queue," "a speech recognition engine [which] compares a current audio context with the context associated with a voice command and causes the browser to perform an action based on the comparison," or the equivalent of any of those features.

Just as Figure 5 of Logan et al. does not anticipate Claim 1, the various portions of the specification of Logan et al. cited by the Examiner do not anticipate Claim 1, either. For example, the Examiner has relied on the following passages to show that Logan et al. discloses "a context-based audio queue ordered based on contents of a page being audibly read by the multi-modal browser to a user" (Office Action at 3):

> As contemplated by the invention, information which is available in text form from news sources, libraries, etc. may be converted to compressed audio form either by human readers or by conventional speech synthesis. If speech synthesis is used, the conversion of text to speech is preferably performed at the client station 103 by the player. In this way, text information alone may be rapidly downloaded from the server 101 since it requires much less data than equivalent compressed audio files, and the downloaded text further provides the user with ready access to a transcript of voice presentations. In other cases, where it is important to capture the quality and authenticity of the original analog speech signals, a text transcript file which collaterally accompanies a compressed voice audio file may be stored in the database 133 from which a transcript may be made available to the user upon request.

(Logan et al., column 5, lines 16-45); as well as

> As hereinafter described in connection with FIG. 5, each voice or text program segment preferably includes a sequencing file which contains the identification of highlighted passages and hypertext anchors within the program content. This sequencing file may further contain references to image files and the start and ending offset locations in the audio presentation when each image display should begin and end. In this way, the image presentation may be synchronized with the audio programming to provide coherent multimedia programming.

(Logan et al., column 5, lines 6-15); and

In addition, the structured program files may advantageously contain, where appropriate, "hyperlink" passages, which may take the form of announced cross references to other materials, or sentences or phrases which describe related information contained elsewhere in the download compilation but which do not follow immediately in the sequence. In order to alert the listener to the fact that a sentence or passage is a hyperlink to other information which is out of the normal playback sequence, an audible cue may advantageously proceed, accompany, or immediately follow the passage in the normal playback which identifies the character of the hyperlinked material. Using the terminology typically employed to described hypertext, the normal programming sequence includes "anchor" passages which are identified by an audible cue signal of some type and are further associated with a reference to hyperlinked material to which the playback may jump upon the listener's request. Hyperlinked material, like all other programming, is advantageously preceded with a topic description and, if the hyperlinked material is a narrative, it should begin with a summary paragraph, followed by increasing detail.

A hyperlink may be directed to a program segment which is not present in the current selections list. In that case, the Link variable contains a negative number to distinguish it from references to a particular Selection_Record, and is interpreted as the negative of a ProgramID number. If the referenced ProgramID is available in the player's mass storage system, it may be fetched an played and, upon its conclusion, an automatic return is made to the original sequence. If the referenced ProgramID does not refer to a locally stored record, the listener is informed that it is currently unavailable, but will be included in the next download for the next session.

> In addition to having means for accepting a user command to execute a jump to the hypertext material, the player also advantageously includes a mechanism (special key or voice command response) which, when activated, causes a "return" to be made to the playing sequence at the point of the original anchor from which the hyperlink was performed. In this way, a listener may listen to as much or as little of the linked information as desired, retaining the ability to return to the original. Just as computer subroutines may be nested by saving the return addresses of a calling instruction in a stack mechanism, a hyperlink may be executed from within a hyperlinked narrative, and so on, with the listener retaining the ability to execute a like

(Logan et al., column 30, lines 20-66) The portions of the disclosure of Logan et al. cited by the Examiner do not refer to a context-based audio queue, especially given the fact that Logan et al. does not address matters involving context such as are addressed by the claimed invention.

Similarly, the Examiner has relied on the following passages to show that Logan et al. discloses "a speech recognition engine for recognizing and registering voice commands, wherein said speech recognition means compares a current audio context with the context associated with a voice command and causes the browser to perform an action based on the comparison." (Office Action at 3):

> The player 103 further includes a sound card 110 which receives audio input from a microphone input device 111 for accepting voice dictation and commands from a user and which delivers audio output to a speaker 113 in order to supply audio information to the user.

(Logan et al., column 3, lines 32-37); as well as

> User Playback Controls
>
> The player mechanism seen at 103 includes both a keyboard and a microphone for accepting keyed or voice commands respectively which

> control the playback mechanism. As indicated at 261, the receipt of a command, which may interrupt the playback of the current selection, and the character of the command is evaluated at 262 to select one of six different types of functions.

(Logan et al., column 12, lines 50-58); and

> Whenever the user issues a "Go" command (seen at 265 in FIG. 3), the player will execute a hyperlink jump to the location indicated by the last "L" record in the selection file. When the jump is made, the location in the "L" record is inserted into the CurrentPlay register 353 after the previous contents of the CurrentPlay register are saved in (pushed into) a zero-based stack 390 at the stack cell location specified by the contents of a StackPtr register 392, which is then incremented. Whenever the listener issues a "Return" command, the previously pushed selection file record location is popped from the stack 390 and returned to the CurrentPlay register 353, and the StackPtr register 392 is decremented. A "Return" command issued when StackPtr=zero (indicating an empty stack) produces no effect.

(Logan et al., column 35, lines 1-15). While the cited portions of the disclosure of Logan et al. contemplate the use of speech recognition as a general matter, there is nothing to anticipate the possibility of context-sensitive uses of speech recognition, which is characteristic of Claim 1.

Applicants respectfully submit that the disclosure of Logan et al. does not anticipate Claim 1 of the claimed invention.

Claim 12. The Examiner has found that "[r]egarding independent claim 12, *Logan et al.* discloses a computer implemented method for controlling an audio player with voice commands" in a manner that anticipates the claimed invention. Applicants respectfully traverse.

Independent Claim 12 of the claimed invention does not make reference to "an audio player" but instead provides as follows:

A computer implemented method for providing context based verbal commands to a multi-modal browser, comprising the steps of:

building a context based audio queue based on the contents of markup language page being audibly read by the multi-modal browser to a user;

storing a current context of the audio queue; and

recognizing and registering voice commands, wherein the current audio context is compared with a voice command, thereby causing the multi-modal browser to perform an action based on the comparison.

(Claim 12, lines 1-8) Thus, the claimed invention deals, to a great extent, with presenting a markup based document both audibly and visually. Like Claim 1, Claim 12 claims context sensitivity, which is absent from the disclosure of Logan et al. While there is mention of the use of natural language text for generating an audio questionnaire in Logan et al. (Logan et al., claim 17), that is only under very specific circumstances not relevant to Claim 12 or other claims of the claimed invention. The disclosure of Logan et al. appears for the most part, to deal with recording and exchanging responses with subscribers, which is not the focus of the claimed invention.

In finding Claim 12 to be anticipated by Logan et al., the Examiner has relied on Figure 5, discussed above, and on Figure 1 from the disclosure of Logan et al.:

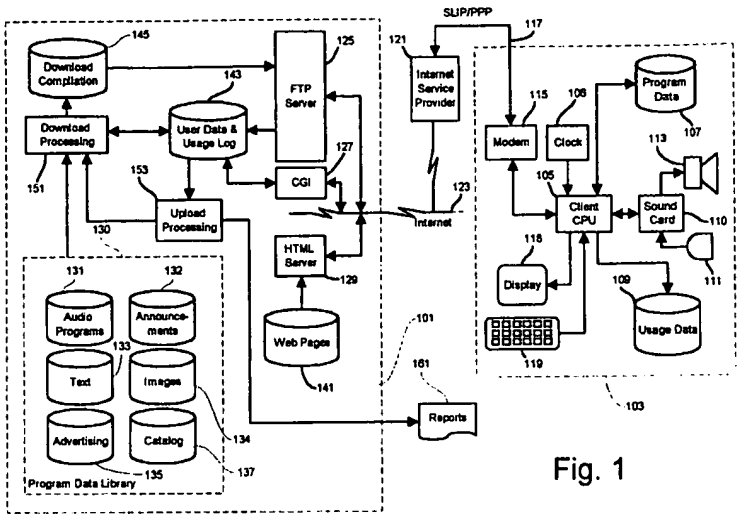


Fig. 1

(Logan et al., Figure 1) Nothing in Figure 1 or 5 of Logan et al. refers to a "computer implemented method for providing context based verbal commands," "building a context based audio queue based on the contents of markup language page being audibly read by the multi-modal browser," "storing a current context of the audio queue," "recognizing and registering voice commands, wherein the current audio context is compared with a voice command," "causing the multi-modal browser to perform an action based on the comparison," or the equivalent of any of those features.

Just as Figures 1 and 5 of Logan et al. do not anticipate Claim 12, the various portions of the specification of Logan et al. cited by the Examiner do not anticipate Claim 12, either. For example, the Examiner has relied on the following passages to show that Logan et al. discloses "building a context-based audio queue based on the contents of markup language page being audibly read by the multi-modal browser to a user" (Office Action at 4):

As contemplated by the invention, information which is available in text form from news sources, libraries, etc. may be converted to compressed audio form either by human readers or by conventional speech synthesis. If speech synthesis is used, the conversion of text to speech is preferably performed at the client station 103 by the player. In this way, text information alone may be rapidly downloaded from the server 101 since it requires much less data than equivalent compressed audio files, and the downloaded text further provides the user with ready access to a transcript of voice presentations. In other cases, where it is important to capture the quality and authenticity of the original analog speech signals, a text transcript file which collaterally accompanies a compressed voice audio file may be stored in the database 133 from which a transcript may be made available to the user upon request.

The host server 101 further stores web page data 141 which is made available to the player 103 by means of the HTML interface 128. The host server 101 additionally stores and maintains a user data and usage log database indicated at 143 which stores uploaded usage data received from the store 109 in the player 103 via the Internet pathway 123 and the FTP server interface 125. The user data 143 further contains additional data describing the preferences, demographic characteristics and program selections unique to each subscriber which is developed largely from user-supplied data obtained when users submit HTML form data via the Internet pathway 123 for processing by the CGI mechanism 127. (Logan et al., column 5, lines 16-45); as well as

As hereinafter described in connection with FIG. 5, each voice or text program segment preferably includes a sequencing file which contains the identification of highlighted passages and hypertext anchors within the program content. This sequencing file may further contain references to

image files and the start and ending offset locations in the audio presentation when each image display should begin and end. In this way, the image presentation may be synchronized with the audio programming to provide coherent multimedia programming.

(Logan et al., column 5, lines 6-15);

Hyperlink Jumps

In addition, the structured program files may advantageously contain, where appropriate, "hyperlink" passages, which may take the form of announced cross references to other materials, or sentences or phrases which describe related information contained elsewhere in the download compilation but which do not follow immediately in the sequence. In order to alert the listener to the fact that a sentence or passage is a hyperlink to other information which is out of the normal playback sequence, an audible cue may advantageously proceed, accompany, or immediately follow the passage in the normal playback which identifies the character of the hyperlinked material. Using the terminology typically employed to described hypertext, the normal programming sequence includes "anchor" passages which are identified by an audible cue signal of some type and are further associated with a reference to hyperlinked material to which the playback may jump upon the listener's request. Hyperlinked material, like all other programming, is advantageously preceded with a topic description and, if the hyperlinked material is a narrative, it should begin with a summary paragraph, followed by increasing detail.

A hyperlink may be directed to a program segment which is not present in the current selections list. In that case, the Link variable contains a negative number to distinguish it from references to a particular Selection_Record, and is interpreted as the negative of a ProgramID number. If the referenced ProgramID is available in the player's mass

storage system, it may be fetched an played and, upon its conclusion, an automatic return is made to the original sequence. If the referenced ProgramID does not refer to a locally stored record, the listener is informed that it is currently unavailable, but will be included in the next download for the next session.

In addition to having means for accepting a user command to execute a jump to the hypertext material, the player also advantageously includes a mechanism (special key or voice command response) which, when activated, causes a "return" to be made to the playing sequence at the point of the original anchor from which the hyperlink was performed. In this way, a listener may listen to as much or as little of the linked information as desired, retaining the ability to return to the original. Just as computer subroutines may be nested by saving the return addresses of a calling instruction in a stack mechanism, a hyperlink may be executed from within a hyperlinked narrative, and so on, with the listener retaining the ability to execute a like

(Logan et al., column 30, lines 19-66);

The Selections File

FIG. 5 shows an illustrative sequence of Selection_Records making up a selection file indicated generally at 351 which illustrates the manner in which the user may navigate the playback session between playback positions designated by the selection file. At any given moment, the next item of programming to be played is specified by an integer register CurrentPlay seen at 353 which holds the record number of the particular Selection_Record in the selections file 351 to be played next. As shown, CurrentPlay points to a subject Selection_Record identified by the LocType "S" 355 and a Location field 357 which contains the ProgramID of an announcement program segment which describes the subject. If the

> user issues a skip command during or shortly after the time when subject announcement is played, the player executes a skip to the next subject, which is accomplished by scanning the selection file 351 until the next subject Selection_Record seen at 360 is located, and then performing a jump by inserting the location of Selection_Record 360 into the CurrentPlay register 353, causing the intervening material to be skipped as indicated by the dashed line 362.

(Logan et al., column 33, lines 29-50);

> The player 103 further includes a sound card 110 which receives audio input from a microphone input device 111 for accepting voice dictation and commands from a user and which delivers audio output to a speaker 113 in order to supply audio information to the user.

(Logan et al., column 3, lines32-37); and

> Whenever the user issues a "Go" command (seen at 265 in FIG. 3), the player will execute a hyperlink jump to the location indicated by the last "L" record in the selection file. When the jump is made, the location in the "L" record is inserted into the CurrentPlay register 353 after the previous contents of the CurrentPlay register are saved in (pushed into) a zero-based stack 390 at the stack cell location specified by the contents of a StackPtr register 392, which is then incremented. Whenever the listener issues a "Return" command, the previously pushed selection file record location is popped from the stack 390 and returned to the CurrentPlay register 353, and the StackPtr register 392 is decremented. A "Return" command issued when StackPtr=zero (indicating an empty stack) produces no effect.

(Logan et al., column 35, lines 1-15) There is thus nothing in the cited portions of Logan et al. which anticipates the use of context sensitivity, either in connection with a multi-modal browser or otherwise.

Applicants respectfully submit that the disclosure of Logan et al. does not anticipate Claim 12 of the claimed invention.

Claims 2 and 13. The Examiner has found that "[r]egarding claims 2 and 13, *Logan et al.* discloses the Program_Segments record URL field specifies the location file containing the program segment in the file storage facility 304 (column 17, line 62 to column 18, line 16; Figure 4); thus, the user listens to audio segments as stored resources based on URL[]s." (Office Action at 5-6) Applicants respectfully traverse.

Because Claim 2 is dependent from Claim 1 and Claim 13 is dependent from Claim 12, Applicants hereby incorporate by reference the foregoing discussion of Claims 1 and 12. Claim 2 of the claimed invention provides as follows:

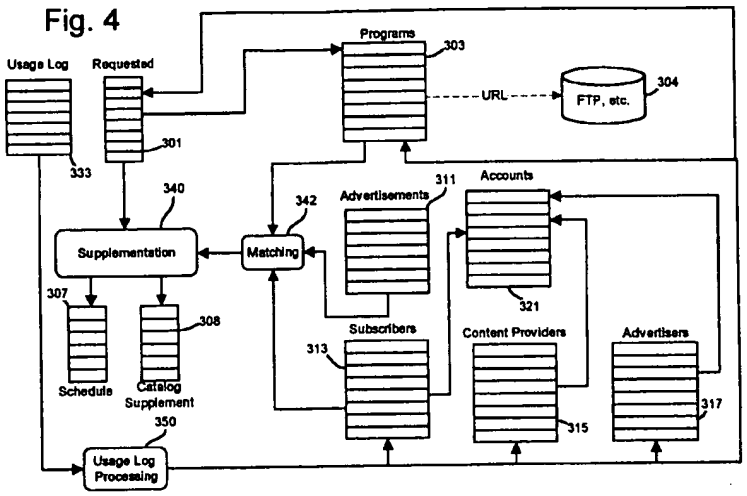
> The system as recited in claim 1, wherein the browser action comprises accessing a different Uniform Resource Locator (URL) and rendering a page specified by the URL.

(Claim 2, lines 1-3) Claim 13 of the claimed invention provides as follows:

> The computer implemented method for providing context based verbal commands to a multi-modal browser as recited in claim 12, wherein the browser action comprises accessing a different Uniform Resource Locator (URL) and displaying the contents of the URL.

(Claim 13, lines 1-4).

In finding Claims 2 and 13 to be anticipated by Logan et al., the Examiner has relied on Figure 4 from the disclosure of Logan et al.:



U.S. Patent Mar. 24, 1998 Sheet 4 of 7 5,732,216

(Logan et al., Figure 4) Figure 4 contemplates locating audio files over the Internet and playing them but does not anticipate "wherein the browser action comprises accessing a different Uniform Resource Locator." In addition, Figure 4 does not require use of a browser as the means to access files over the Internet.

Just as Figure 4 of Logan et al. does not anticipate Claim 2 or Claim 13 of the claimed invention, the portion of the specification of Logan et al. cited by the Examiner do not anticipate the claims, either. The Examiner has relied on the following passages to show that Logan et al. anticipates Claims 2 and 13:

> The Program_Segment record's URL field specifies the location of the file containing the program segment in the file storage facility indicated at 304 in FIG. 4 (i.e., normally on the FTP server 125 seen in FIG. 1, but potentially including storage areas on the web server 141 or at any other accessible location on the Internet). In addition, the subscriber

> may wish to designate for future play a program segment already loaded into the player 103 by virtue of a prior download. The subscriber may elect to include an already loaded file because it was not reached in a prior playback session or because the subscriber wishes replay the selection. In that event, the ProgramID of such a selection is nonetheless included in the uploaded selection list (Requested Table 301), recognizing that at the time of actual download, the player 103 will only request the transfer of those program segments not already present in local storage. The uploaded Requested list 301 should accordingly be understood to be indicative of the requested content of a future planned playback session and not necessarily a listing of programs to be downloaded. The selection of files to download is preferably made by the player which issues FTP download requests from the server by specifying the URLs of the needed files.

(Logan et al., column 17, line 62 – column 18, line 16) While the cited passage may show what the Examiner describes it as showing, it nonetheless does not anticipate Claim 2 or Claim 13 because it does not disclose the substance Claim 1 while adding "wherein the browser action comprises accessing a different Uniform Resource Locator (URL) and rendering a page specified by the URL" (Claim 2) and because it does not disclose the substance of Claim 12 while adding "wherein the browser action comprises accessing a different Uniform Resource Locator (URL) and displaying the contents of the URL." (Claim 13) Thus, while Claims 2 and 13 do claim use of a URL, the substance of the claims is not anticipated by the portion of the disclosure of Logan et al. cited by the Examiner in support of rejection.

Applicants respectfully submit that Claims 2 and 13 of the claimed invention are not anticipated by the disclosure of Logan et al.

Claim 11. The Examiner has found that "[r]egarding claim 11, *Logan et al.* discloses the host server stores web page data 141 by means of an HTNL interface . . . HTML web server 129 presents HTML program selection forms . . . narrative text is

presented in the interactive, multimedia format expressed in the first instance using essentially in the interactive, multimedia format expressed in the finrst instance using essentially conventional hypertext markup language." (Office Action at 6)

Applicants respectfully traverse the rejection of Claim 11. Because Claim 11 is dependent from Claim 1, Applicants hereby incorporate by reference the foregoing discussion of Claim 1. Claim 11 of the claimed invention provides as follows:

> The system as recited in claim 1, wherein the page being audibly read is a markup language page.

(Claim 11, lines 1-2) Thus, the Examiner has found that the disclosure of Logan et al. anticipates that "a page being read by [a] multi-modal browser to a user" (Claim 1, lines 3-4) may be "a markup language page." (Claim 11, line 2) Because Claim 1 is not anticipated by Logan et al., as discussed above, there does not appear to be a basis for concluding that Claim 11 is anticipated by Logan et al.

In finding Claim 11 to be anticipated by Logan et al., the Examiner has relied on Figure 1, set forth above, and Figure 7 from the disclosure of Logan et al.:

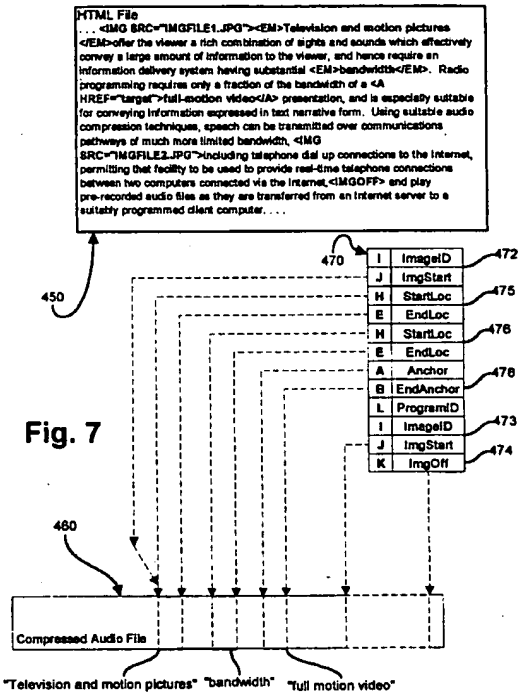


Fig. 7

(Logan et al., Figure 7) Nothing in Figure 1 or 7 of Logan et al. discloses the substance Claim 1, including context based features, while adding "wherein the page being audibly read is a markup language page." (Claim 11)

Just as Figures 1 and 7 of Logan et al. do not anticipate Claim 11, the various portions of the specification of Logan et al. cited by the Examiner also do not anticipate Claim 11. The Examiner has relied on the following passages to show that Logan et al. discloses "the host server stores web page data 141 by means of an HTML interface." (Office Action at 6):

> The host server 101 further stores web page data 141 which is made available to the player 103 by means of the HTML interface 128. The host server 101 additionally stores and maintains a user data and usage log database indicated

(Logan et al., column 5, lines 32-35) While the cited passage may show what the Examiner describes it as showing, it nonetheless does not anticipate Claim 11 because it does not disclose the substance Claim 1, including context based features, while adding "wherein the page being audibly read is a markup language page." (Claim 11)

In addition, the Examiner has relied on the following portion of the disclosure of Logan et al. to show that "HTML web server 129 presents HTML program selection forms." (Office Action at 6):

> In addition to the downloaded catalog of available items which may be viewed by the subscriber from the available downloaded information, the user may re-establish an Internet connection to the HTML web server 129 which presents HTML program selection and search request forms, enabling the subscriber to locate remotely stored programming which may be of particular interest to the subscriber. When such programs are selected in the HTML session, the user's additional preferences and selections may be posted into the user data file 143 and the identification of the needed files may be passed to the client/player 103 for inclusion in the next download request.

(Logan et al., column 8, lines 48-60) Again, while the cited passage may show what the Examiner describes it as showing, it nonetheless does not anticipate Claim 11 because it does not disclose the substance Claim 1, including context based features, while adding "wherein the page being audibly read is a markup language page." (Claim 11)

Finally, the Examiner has relied on the following portion of the disclosure of Logan et al. to show that "narrative text is presented in the interactive, multimedia format expressed in the first instance using essentially conventional hypertext markup language."

(Office Action at 6):

the usage log is transferred (see 219, FIG. 2).

Defining Audio Programming with HTML

Narrative text to be presented in the interactive, multimedia format made possible by the present invention may be advantageously expressed in the first instance using essentially conventional hypertext markup language, "HTML". FIG. 7 shows an example of the content of a portion of an illustrative HTML text file indicated generally at 450 used to create an audio file seen at 460 and a selections file indicated at 470.

The HTML file illustrated at 450 uses conventional <IMG> tags to identify image files, conventional emphasizing tag pairs <EM> and </EM> to designate highlighted passages, and conventional <A> and </A> HTML tag pairs to designate the anchor text and link target of a hypertext link. Utilizing conventional HTML to describe the narrative content to be presented in audio form provides several significant advantages, not the least of which are:

conventional HTML composition software may be used to add the image and emphasis tags by means of visual tools which eliminate the need for hand-coding on a character level;

(a) a narrative text version of the audio programming may be viewed and printed, including both the emphasized text and the imbedded images, using most popular web browsers;

existing HTML files may be readily converted into audio multimedia presentations with little or no HTML editing being required;

HTML file may be made available from a server in a form which can be viewed in the normal way by any web browser yet

> and alternatively presented accordance with the invention in the form of an interactively browsable audio program with synchronized images;
>
> the HTML file may be supplied along with the audio file as a transcript for the audio presentation, and to permit the audio presentation to be indexed and searched; and
>
> the HTML may be automatically converted into the combination of an audio file using conventional speech synthesis techniques to process the narrative text with the HTML tags being used to compile a selections file which enables the player to interactively browse the audio file using highlighted and linked passages, and to synchronize the image presentation with the audio file.

(Logan et al., column 43, lines 15-60) Once more, while the cited passage may show what the Examiner describes it as showing, it nonetheless does not anticipate Claim 11 because it does not disclose the substance Claim 1, including context based features, while adding "wherein the page being audibly read is a markup language page." (Claim 11)

Applicants respectfully submit that Claim 11 of the claimed invention is not anticipated by the disclosure of Logan et al.

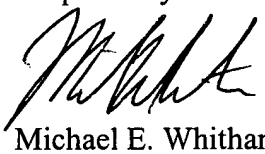
Conclusion

In view of the foregoing, it is respectfully requested that the application be reconsidered, that Claims 1-21 be allowed, and that the application be passed to issue.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary in a telephonic or personal interview.

A provisional petition is hereby made for any extension of time necessary for the continued pendency during the life of this application. Please charge any fees for such provisional petition and any deficiencies in fees and credit any overpayment of fees to Deposit Account No. 09-0457 (IBM-Endicott).

Respectfully submitted,



Michael E. Whitham
Registration No.32,635

Whitham, Curtis & Christofferson, P.C.
11491 Sunset Hills Road, Suite 340
Reston, VA 20190
Tel. (703) 787-9400
Fax. (703) 787-7557